Land Information System (LIS) Test Plan

Submitted under Task Agreement GSFC-CT-2

Cooperative Agreement Notice (CAN)
CAN-00OES-01

Increasing Interoperability and Performance of Grand Challenge Applications in the Earth, Space, Life, and Microgravity Sciences

January 7, 2003

Draft
Revision 1.0

History:

Revision	Summary of Changes	Date
1.0	Draft	January 7, 2003

1 Scope

Provide overview information about the project and the software test plan. The following sub-sections provide an overview of the LIS project.

1.1 LIS

The Land Information System (LIS) will have the following components: (1) A high-resolution (1km) Land Data Assimilation System (LDAS), involving several independent community Land Surface Models (LSMs), land surface data assimilation technologies, and integrated database operations for observation and prediction data management; and (2) A web-based user interface based on the GRid Analysis and Display System (GrADS) and the Distributed Oceanographic Data System (DODS) for accessing data mining, numerical modeling, and visualization tools. The LIS will be available as a "production" system on a centralized server for large applications. By incorporating and promulgating the existing Assistance for Land_surface Modeling Activities (ALMA; http://www.lmd.jussieu.fr/ALMA/) and DODS standards for model coupling and data visualization, LIS will contribute to the definition of the land surface modeling and assimilation standards for the Earth System Modeling Framework (ESMF).

1.2 Land Surface Modeling and Data Assimilation

In general, land surface modeling seeks to predict the terrestrial water, energy and biogeochemical processes by solving the governing equations of the soil-vegetation-snowpack medium. Land surface data assimilation seeks to synthesize data and land surface models to improve our ability to predict and understand these processes. The ability to predict terrestrial water, energy and biogeochemical processes is critical for applications in weather and climate prediction, agricultural forecasting, water resources management, hazard mitigation and mobility assessment.

In order to predict water, energy and biogeochemical processes using (typically 1-D vertical) partial differential equations, land surface models require three types of inputs: 1) initial conditions, which describe the initial state of land surface; 2) boundary conditions, which describe both the upper (atmospheric) fluxes or states also known as "forcings" and the lower (soil) fluxes or states; and 3) parameters, which are a function of soil, vegetation, topography, etc., and are used to solve the governing equations.

1.3 Land Data Assimilation System (LDAS)

LDAS is a model control and input/output system (consisting of a number of subroutines, modules written in Fortran 90 source code) that drives multiple offline one dimensional land surface models (LSMs) using a vegetation defined "tile" or "patch" approach to simulate sub-grid scale variability. The one-dimensional LSMs such as CLM and NOAH, which are subroutines of LDAS, apply the governing equations of the physical processes of the soil-vegetation-snowpack medium. These land surface models aim to characterize the transfer of mass, energy, and momentum between a vegetated surface and the atmosphere.

LDAS makes use of various satellite and ground based observation systems within a land data assimilation framework to produce optimal output fields of land surface states and fluxes. The LSM predictions are greatly improved through the use of a data assimilation environment such as the one provided by LDAS. In addition to being forced with real time output from numerical prediction models and satellite and radar precipitation measurements, LDAS derives model parameters from existing topography, vegetation and soil coverages. The model results are aggregated to various temporal and spatial scales, e.g., 3 hourly, $1/4°$.

The execution of LDAS starts with reading in the user specifications. The user selects the model domain and spatial resolution, the duration and timestep of the run, the land surface model, the type of forcing from a list of model and observation based data sources, the number of "tiles" per grid square (described below), the soil parameterization scheme, reading and writing of restart files, output specifications, and the functioning of several other enhancements including elevation correction and data assimilation.

The system then reads the vegetation information and assigns subgrid tiles on which to run the one-dimensional simulations. LDAS runs its 1-D land models on vegetation-based "tiles" to simulate variability below the scale of the model grid squares. A tile is not tied to a specific location within the grid square. Each tile represents the area covered by a given vegetation type.

Memory is dynamically allocated to the global variables, many of which exist within Fortran 90 modules. The model parameters are read and computed next. The time loop begins and forcing data is read, time/space interpolation is computed and modified as necessary. Forcing data is used to specify boundary conditions to the land surface model. The LSMs in LDAS are driven by atmospheric forcing data such as precipitation, radiation, wind speed, temperature, humidity, etc., from various sources. LDAS applies spatial interpolation to convert forcing data to the appropriate resolution required by the model. Since the forcing data is read in at certain regular intervals, LDAS also temporally interpolates time average or instantaneous data to that needed by the model at the current timestep. The selected model is run for a vector of "tiles", intermediate information is stored in modular arrays, and output and restart files are written at the specified output interval.

1.4 Community Land Model (CLM)

CLM is a 1-D land surface model, written in Fortran 90, developed by a grass-roots collaboration of scientists who have an interest in making a general land model available for public use. LDAS currently uses CLM version 1.0, formerly known as common land model. CLM version 2.0 was released in May 2002 and will be implemented in LDAS in future. The source code for CLM 2.0 is freely available from the National Center for Atmospheric Research (NCAR) (`http://www.cgd.ucar.edu/tss/clm/`). The CLM is used as the land model for the community climate system model (CCSM) (`http://www.ccsm.ucar.edu/`) and the community atmosphere model (CAM) (`http://www.cgd.ucar.edu/cms/`). CLM is executed with all forcing, parameters, dimensioning, output routines, and coupling performed by an external driver of the user's design (in this case done by LDAS). CLM requires pre-processed data such as the land surface type, soil and vegetation parameters, model initialization, and atmospheric boundary conditions as input. The model applies finite-difference spatial discretization methods and a fully implicit time-integration scheme to numerically integrate the governing equations. The model subroutines apply the governing equations of the physical processes of the soil-vegetation-snowpack medium, including the surface energy balance equation, Richards' [4] equation for soil hydraulics, the diffusion equation for soil heat transfer, the energy-mass balance equation for the snowpack, and the Collatz et al. [2] formulation for the conductance of canopy transpiration.

1.5 The Community NOAH Land Surface Model

The community NOAH Land Surface Model is a stand-alone, uncoupled, 1-D column model freely available at the National Centers for Environmental Prediction (NCEP; `ftp://ftp.ncep.noaa.gov/pub/gcp/ldas/noahlsm/`). NOAH can be executed in either coupled or uncoupled mode. It has been coupled with the operational NCEP mesoscale Eta model [1] and its companion Eta Data Assimilation System (EDAS) [5], and the NCEP global Medium-Range Forecast model (MRF) and its companion Global Data Assimilation System (GDAS). When NOAH is executed in uncoupled mode, near-surface atmospheric forcing data (e.g., precipitation, radiation, wind speed, temperature, humidity) is required as input. NOAH simulates soil moisture (both liquid and frozen), soil temperature, skin temperature, snowpack depth, snowpack water equivalent, canopy water content, and the energy flux and water flux terms of the surface energy balance and surface water balance. The model applies finite-difference spatial discretization methods and a Crank-Nicholson time-integration scheme to numerically integrate the governing equations of the physical processes of the soil vegetation-snowpack medium, including the surface energy balance equation, Richards' [4] equation for soil hydraulics, the diffusion equation for soil heat transfer, the energy-mass balance equation for the snowpack, and the Jarvis [3] equation for the conductance of canopy transpiration.

1.6 Variable Infiltration Capacity Model

Variable Infiltration Capacity (VIC) model is a macroscale hydrologic model, written in C, being developed at the University of Washington, and Princeton University. The VIC code repository along with the model description and source code documentation is available `http://www.hydro.washington.edu/Lettenmaier/Models/VIC/VIChome.html`. VIC is used in macroscopic land use models such as SEA - BASINS (`http://boto.ocean.washington.edu/seasia/intro.htm`). VIC is a semi-distributed, grid-based hydrological model, which parameterizes the dominant hydrometeorological processes taking place at the land surface - atmospheric interface. The execution of VIC model requires preprocessed data such as precipitation, temperature, meteorological forcing, soil and vegetation parameters, etc. as input. The model uses three soil layers and one vegetation layer with energy and moisture fluxes exchanged between the layers. The VIC model represents surface and subsurface hydrologic processes on a spatially distributed (grid cell) basis. Partitioning grid cell areas to different vegetation classes can approximate sub-grid scale variation in vegetation characteristics. VIC models the processes governing the flux and storage of water and heat in each cell-sized system of vegetation and soil structure. The water balance portion of VIC is based on three concepts:
1) Division of grid-cell into fraction sub-grid vegetation coverage.
2) The variable infiltration curve for rainfall/runoff partitioning at the land surface.
3) A baseflow/deep soil moisture curve for lateral baseflow.

Water balance calculations are preformed at three soil layers and within a vegetation canopy. An energy balance is calculated at the land surface. A full description of algorithms in VIC can be found in the references listed at the VIC website.

1.7 Software test plan overview

This sub-section provides an overview of the software test plan.

1.8 System Overview

Provide a brief description of the software system to be tested.

2 Referenced Documents

List other project documentation that that this plan refers to. Such as:

- Software Development Plan
- Requirements Document
- Software Design Document

3 Software Test Environment

Describe the test environment(s). Include environment diagrams. This section should refer to:

- Software system version or build being tested
- Any simulators
- Drivers and test tools
- Compilers and libraries
- Input data sets
- Hardware platforms (SGI Origin 3000, Beowulf cluster, etc)
- Participating organizations and personnel
- Briefly describe the process for configuring the test environment(s):

3.1 Software build being tested

- LIS_MILESTONE_F_TESTING
- LIS_MILESTONE_I_TESTING

3.2 Drivers and test tools

- Customized PBS batch queuing script
- SpeedShop performance tools

3.3 Compilers and libraries

- Compilers: f90, c
- Libraries: mpi, hdf, w3lib, bacio, iplib, gfio

3.4 Input data sets

- BCS data
- GVEG data
- GEOS data
- NRL data
- AGRMET data

See Appendix for a full listing of the input data files.

3.5 Hardware platforms

- SGI Origin 3000 – see http://www.nas.nasa.gov/User/Systemsdocs/O3K/o3k.html
- LIS Linux Cluster – see LIS Design Document

3.6 Configuration of the test environment(s)

- Checkout appropriate build from CVS repository.
- Edit namelist parameters in ldas.crd file (henceforth referred to as the card file).
- Edit the PBS batch queuing script.

4 Test Identification

This section describes the tests to be performed under this plan.

4.1 General Information

This section contains general information applicable to the overall testing to be performed.

4.1.1 Test Level

Specify the level at which testing will be performed (component testing, integration testing, system testing).
Testing will be performed at the system level.

4.1.2 Test Classes

Specify the type of tests that will be performed (performance, invalid input, capacity tests, operational tests).
The tests to be performed will be performance and operational tests.

4.2 Planned Tests

This section describes the specific tests to be performed under this plan.

4.2.1 Example Test #1

For each test, include the following information:

- Purpose of the test - Summary of capabilities to be verified
- Test Type/Class - performance, functional, etc.
- Test inputs - data files, databases, user inputs, configuration parameters, etc.

- Verification method(s) - as defined in the associated requirements specification(s)
- Special Requirements - (e.g., 48 hours of continuous machine time, etc.)
- Assumptions/Constraints
- Expected results - description of expected outcome
- Actual results (added during the testing phase) - description of observed results in comparison to expected results
- Reference any discrepancy reports that were written.

4.2.2 5km Input Data Validation

- Purpose: To validate the newly generated 5km input data sets.
- Type/Class: ?
- Test inputs: Raw input data sets ***LIST***
- Verification methods: Generated input data shall be reviewed by scientist.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Generated input data shall be correct.
- Actual results:
- Discrepancy reports:

4.2.3 5km CLM Run on SGI Origin

- Purpose: To verify that LDAS with CLM will run on the SGI Origin at a 5km resolution.
- Type/Class: Operational
- Test inputs: 5km input data sets ***LIST***, ldas card file
- Verification methods: Generated output data shall be reviewed by scientist.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Generated output data shall be correct.
- Actual results:
- Discrepancy reports:

4.2.4 5km NOAH Run on SGI Origin

- Purpose: To verify that LDAS with NOAH will run on the SGI Origin at a 5km resolution.
- Type/Class: Operational
- Test inputs: 5km input data sets ***LIST***, ldas card file
- Verification methods: Generated output data shall be reviewed by scientist.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Generated output data shall be correct.
- Actual results:
- Discrepancy reports:

4.2.5 5km VIC Run on SGI Origin

- Purpose: To verify that LDAS with VIC will run on the SGI Origin at a 5km resolution.
- Type/Class: Operational
- Test inputs: 5km input data sets ***LIST***, ldas card file
- Verification methods: Generated output data shall be reviewed by scientist.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Generated output data shall be correct.
- Actual results:
- Discrepancy reports:

4.2.6 5km CLM Run on LIS Linux Cluster

- Purpose: To verify that LDAS with CLM will run on the LIS Linux Cluster at a 5km resolution.
- Type/Class: Operational
- Test inputs: 5km input data sets ***LIST***, ldas card file

- Verification methods: Generated output data shall be reviewed by scientist.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Generated output data shall be correct.
- Actual results:
- Discrepancy reports:

4.2.7 5km NOAH Run on LIS Linux Cluster

- Purpose: To verify that LDAS with NOAH will run on the LIS Linux Cluster at a 5km resolution.
- Type/Class: Operational
- Test inputs: 5km input data sets ***LIST***, ldas card file
- Verification methods: Generated output data shall be reviewed by scientist.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Generated output data shall be correct.
- Actual results:
- Discrepancy reports:

4.2.8 5km VIC Run on LIS Linux Cluster

- Purpose: To verify that LDAS with VIC will run on the LIS Linux Cluster at a 5km resolution.
- Type/Class: Operational
- Test inputs: 5km input data sets ***LIST***, ldas card file
- Verification methods: Generated output data shall be reviewed by scientist.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Generated output data shall be correct.
- Actual results:
- Discrepancy reports:

4.2.9 5km CLM Run on SGI Origin

- Purpose: To verify that LDAS with CLM will run on the SGI Origin at a 5km resolution at a throughput of 1 ms per grid cell per day.
- Type/Class: Performance
- Test inputs: 5km input data sets ***LIST***, ldas card file
- Verification methods: Throughput value will be calculated from run-time value.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Throughput will be less than or equal to 1 ms per grid cell per day.
- Actual results:
- Discrepancy reports:

4.2.10 5km NOAH Run on SGI Origin

- Purpose: To verify that LDAS with NOAH will run on the SGI Origin at a 5km resolution at a throughput of 1 ms per grid cell per day.
- Type/Class: Performance
- Test inputs: 5km input data sets ***LIST***, ldas card file
- Verification methods: Throughput value will be calculated from run-time value.
- Special requirements: ?
- Assumptions/Constraints: ?
- Expected results: Throughput will be less than or equal to 1 ms per grid cell per day.
- Actual results:
- Discrepancy reports:

5 Test Schedules

Provide a schedule of the testing that will be performed for the project and the milestone in which the testing will occur. This might include build testing, integration testing, system level testing, acceptance testing, regression testing.

Test	Date	Milestone
4.2.2	Jan 2003	-
4.2.3	Feb 2003	F
4.2.4	Feb 2003	F
4.2.9	Feb 2003	F
4.2.10	Feb 2003	F
4.2.5	Jun 2003	I
4.2.6	Jun 2003	I
4.2.7	Jun 2003	I
4.2.8	Jun 2003	I

6 Requirements Traceability

Provide an updated traceability matrix which maps the Test name or Identifier to the requirements addressed by that test.

References

[1] F. Chen, K. Mitchell, J. Schaake, Y. Xue, H. Pan, V. Koren, Y. Duan, M. Ek, and A. Betts. Modeling of land-surface evaporation by four schemes and comparison with fife observations. *J. Geophys. Res.*, 101(D3):7251–7268, 1996.

[2] G. J. Collatz, C. Grivet, J. T. Ball, and J. A. Berry. Physiological and environmental regulation of stomatal conducatance: Photosynthesis and transpiration: A model that includes a laminar boundary layer. *Agric. For. Meteorol.*, 5:107–136, 1991.

[3] P. G. Jarvis. The interpretation of leaf water potential and stomatal conductance found in canopies of the field. *Phil. Trans. R. Soc.*, B(273):593–610, 1976.

[4] L. A. Richards. Capillary conduction of liquids in porous media. *Physics*, 1:318–333, 1931.

[5] E. Rogers, T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. Baldwin, N. W. Junker, and Y. Lin. Changes to the operational "early" eta analysis/forecast system at the national centers of environmental prediction. *Wea. Forecasting*, 11:391–413, 1996.